

Parameterized Complexity of MaxSat Above Average*

Robert Crowston[†] Gregory Gutin[†] Mark Jones[†]
 Venkatesh Raman[‡] Saket Saurabh[‡]

This paper is dedicated to the memory of Alan Turing

Abstract

In MAXSAT, we are given a CNF formula F with n variables and m clauses and asked to find a truth assignment satisfying the maximum number of clauses. Let r_1, \dots, r_m be the number of literals in the clauses of F . Then $\text{asat}(F) = \sum_{i=1}^m (1 - 2^{-r_i})$ is the expected number of clauses satisfied by a random truth assignment (the truth values to the variables are distributed uniformly and independently). It is well-known that, in polynomial time, one can find a truth assignment satisfying at least $\text{asat}(F)$ clauses. In the parameterized problem MAXSAT-AA, we are to decide whether there is a truth assignment satisfying at least $\text{asat}(F) + k$ clauses, where k is the (nonnegative) parameter. We prove that MAXSAT-AA is para-NP-complete and thus, MAXSAT-AA is not fixed-parameter tractable unless $P=NP$. This is in sharp contrast to the similar problem MAXLIN2-AA which was recently proved to be fixed-parameter tractable by Crowston *et al.* (FSTTCS 2011).

In fact, we consider a more refined version of MAXSAT-AA, MAX- $r(n)$ -SAT-AA, where $r_j \leq r(n)$ for each j . Alon *et al.* (SODA 2010) proved that if $r = r(n)$ is a constant, then MAX- r -SAT-AA is fixed-parameter tractable. We prove that MAX- $r(n)$ -SAT-AA is para-NP-complete for $r(n) = \lceil \log n \rceil$. We also prove that assuming the exponential time hypothesis, MAX- $r(n)$ -SAT-AA is not in XP already for any $r(n) \geq \log \log n + \phi(n)$, where $\phi(n)$ is any unbounded strictly increasing function. This lower bound on $r(n)$ cannot be decreased much further as we prove that MAX- $r(n)$ -SAT-AA is (i) in XP for any $r(n) \leq \log \log n - \log \log \log n$ and (ii) fixed-parameter tractable for any $r(n) \leq \log \log n - \log \log \log n - \phi(n)$, where $\phi(n)$ is any unbounded strictly increasing function. The proof uses some results on MAXLIN2-AA.

1 Introduction

SATISFIABILITY is a well-known fundamental problem in Computer Science. Its optimization version (finding the maximum number of clauses that can be satisfied by a truth assignment) and its generalizations (constraint satisfaction problems) are well

*A preliminary version of this paper will appear in the proceedings of LATIN 2012

[†]Royal Holloway, University of London, Egham, Surrey, UK

[‡]The Institute of Mathematical Sciences, Chennai 600 113, India

studied in almost every paradigm of algorithms and complexity including approximation and parameterized complexity. Here we consider the parameterized complexity of a variation of MAXSAT.

In parameterized complexity, one identifies a natural parameter k in the input and algorithms are designed and analyzed to confine the combinatorial explosion to this parameter, while keeping the rest of the running time to be polynomial in the size of the input. More specifically, the notion of feasibility is *fixed-parameter tractability* where one is interested in an algorithm whose running time is $O(f(k)n^c)$, where f is an arbitrary (typically exponential) function of k , n is the input size and c is a constant independent of k . When the values of k are relatively small, fixed-parameter tractability implies that the problem under consideration is tractable, in a sense. The class of fixed-parameter tractable problems will be denoted by FPT. When the available running time is replaced by the much more powerful $n^{O(f(k))}$, we obtain the class XP, where each problem is polynomial-time solvable for any fixed value of k . It is well-known that FPT is a proper subset of XP. More details on parameterized algorithms and complexity are given at the end of this section.

A well-studied parameter in most optimization problems is the size of the solution. In particular, for MAXSAT, the natural parameterized question is whether a given boolean formula in CNF has an assignment satisfying at least k clauses. Using the (folklore) observation that every CNF formula on m clauses has an assignment satisfying at least $m/2$ clauses (a random assignment will satisfy at least $m/2$ clauses), Mahajan and Raman [18] observed that this problem is fixed-parameter tractable. This lower bound, of $m/2$, for the maximum number of clauses, means that the problem is interesting only when $k > m/2$, i.e., when the values of k are relatively large. Hence Mahajan and Raman introduced and showed fixed-parameter tractable, a more natural parameterized question, namely whether the given CNF formula has an assignment satisfying at least $m/2 + k$ clauses.

This idea of parameterizing above a (tight) lower bound has been followed up in many directions subsequently. For MAXSAT alone, better (larger than $m/2$) lower bounds for certain classes of instances (formulas with no pair of unit clauses in conflict, for example) have been proved and the problems parameterized above these bounds have been shown to be fixed-parameter tractable [4, 13]. When every clause has r literals, the expected number of clauses that can be satisfied by a (uniformly) random assignment can easily be seen to be $(1 - 1/2^r)m$, and Alon *et al.* [1] proved that checking whether k more than this many clauses can be satisfied in such an r -CNF formula is fixed-parameter tractable. This problem is known to be MAX- r -SAT-AA. The problem MAX- $r(n)$ -SAT-AA we consider in this paper, is a refinement of this problem, where r need not be a constant.

More specifically, the problem MAXSAT-AA¹, we address is the following.

MAXSAT-AA

Instance: A CNF formula F with clauses c_1, \dots, c_m , and variables x_1, \dots, x_n , and a nonnegative integer k . Clause c_i has r_i literals, $i = 1, \dots, m$.

Parameter: k .

Question: Decide whether there is a truth assignment satisfying at least $\text{asat}(F) + k$ clauses, where $\text{asat}(F) = \sum_{i=1}^m (1 - 2^{-r_i})$.

¹In this paper, AA is an abbreviation for Above Average

The problem MAX- $r(n)$ -SAT-AA is a refinement of MAXSAT-AA in which each clause has at most $r(n)$ literals.

Note that $\text{asat}(F)$ is the average number of satisfied clauses. Indeed, if we assign TRUE or FALSE to each x_j with probability $1/2$ independently of the other variables, then the probability of c_i being satisfied is $1 - 2^{-r_i}$, and by linearity of expectation, $\text{asat}(F)$ is the expected number of satisfied clauses. (Since our distribution is uniform, $\text{asat}(F)$ is indeed the average number of satisfied clauses.) Let $\text{sat}(F)$ denote the maximum number of clauses satisfied by a truth assignment. For Boolean variables y_1, \dots, y_t , the *complete set of clauses on y_1, \dots, y_t* is the set $\{(z_1 \vee \dots \vee z_t) : z_i \in \{y_i, \bar{y}_i\}, i \in [t]\}$. Any formula F consisting of one or more complete sets of clauses shows that the lower bound $\text{sat}(F) \geq \text{asat}(F)$ on $\text{sat}(F)$ is tight. Using the derandomization method of conditional expectations (see, e.g., [2]), it is easy to obtain a polynomial time algorithm which finds a truth assignment satisfying at least $\text{asat}(F)$ clauses. Thus, the question asked in MAXSAT-AA is whether we can find a better truth assignment efficiently from the parameterized complexity point of view.

New Results. Solving an open problem of [6] we show that MAXSAT-AA is not fixed-parameter tractable unless $P=NP$. More specifically, we show this for MAX- $r(n)$ -SAT-AA for $r(n) = \lceil \log n \rceil$. Also, we prove that unless the exponential time hypothesis (ETH) is false, MAX- $r(n)$ -SAT-AA is not even in XP for any $r(n) \geq \log \log n + \phi(n)$, where $\phi(n)$ is any unbounded strictly increasing function². These two results are proved in Section 2.

These results are in sharp contrast to the related problems MAXLIN2-AA (see Section 3 for a definition of this problem) and MAX- r -SAT-AA which are known to be fixed-parameter tractable. Also this is one of the very few problems in the ‘above guarantee’ parameterization world, which is known to be hard. See [19] for a few other hard above guarantee problems.

Then, complementing our hardness results, we show that the lower bound above on $r(n)$ cannot be decreased much further as we prove that MAX- $r(n)$ -SAT-AA is in XP for any $r(n) \leq \log \log n - \log \log \log n$ and fixed-parameter tractable for any $r(n) \leq \log \log n - \log \log \log n - \phi(n)$, where $\phi(n)$ is any unbounded strictly increasing function. This result generalizes the one of Alon *et al.* [1] and is proved in Section 3.

The problem we study is one of the few problems in the ‘above guarantee parameterization’ where we parameterize above an instance-specific bound, as opposed to a generic bound, see [19] for a discussion on this issue. Another example of such parameterizations is the problem VERTEX COVER parameterized above the maximum matching of the given graph. See [7, 21] for recent results on this problem.

We complete this paper in Section 4 by stating an open problem on permutation constraint satisfaction problems parameterized above average.

Basics on Parameterized Complexity. A parameterized problem Π can be considered as a set of pairs (I, k) where I is the *problem instance* and k (usually a nonnegative integer) is the *parameter*. Π is called *fixed-parameter tractable (fpt)* if membership of (I, k) in Π can be decided by an algorithm of runtime $O(f(k)|I|^c)$, where $|I|$ is the size of I , $f(k)$ is an arbitrary function of the parameter k only, and c

²A function f is *strictly increasing* if for every pair x', x'' of values of the argument with $x' < x''$, we have $f(x') < f(x'')$.

is a constant independent from k and I . Such an algorithm is called an *fpt* algorithm. Let Π and Π' be parameterized problems with parameters k and k' , respectively. An *fpt-reduction* R from Π to Π' is a many-to-one transformation from Π to Π' , such that (i) $(I, k) \in \Pi$ if and only if $(I', k') \in \Pi'$ with $k' \leq g(k)$ for a fixed function g , and (ii) R is of complexity $O(f(k)|I|^c)$.

Π belongs to XP if membership of (I, k) in Π can be decided by an algorithm of runtime $|I|^{O(f(k))}$, where $f(k)$ is an arbitrary function of the parameter k only. It is well-known that FPT is a proper subset of XP [8, 9, 20].

Π is in *para-NP* if membership of (I, k) in Π can be decided in nondeterministic time $O(f(k)|I|^c)$, where $|I|$ is the size of I , $f(k)$ is an arbitrary function of the parameter k only, and c is a constant independent from k and I . Here, nondeterministic time means that we can use nondeterministic Turing machine. A parameterized problem Π' is *para-NP-complete* if it is in para-NP and for any parameterized problem Π in para-NP there is an fpt-reduction from Π to Π' . It is well-known that a parameterized problem Π belonging to para-NP is para-NP-complete if we can reduce an NP-complete problem to the subproblem of Π when the parameter is equal to some constant [9].

For example, consider the k -COLORABILITY problem, where given a graph G and a positive integer k (k is the parameter), we are to decide whether G is k -colorable. Since the (unparameterized) COLORABILITY problem is in NP, k -COLORABILITY is in para-NP. k -COLORABILITY is para-NP-complete since 3-COLORABILITY is NP-complete.

For further background and terminology on parameterized algorithms and complexity we refer the reader to the monographs [8, 9, 20].

For an integer n , $[n]$ stands for $\{1, \dots, n\}$.

2 Hardness Results

In this section we give our hardness results. For our results we need the following problem as a starting point for our reductions.

LINEAR-3-SAT

Instance: A 3-CNF formula F with clauses c_1, \dots, c_m , and variables x_1, \dots, x_n such that $m \leq cn$ for some fixed constant c . That is, number of clauses in F is linear in the number of variables.

Question: Decide whether there is a truth assignment satisfying F .

It is well known that LINEAR-3-SAT is NP-complete. For example, the well-known theorem of Tovey [22] states that the 3-SAT problem is NP-complete even when the input consists of 3-CNF formula with every variable contained in at most four clauses.

Theorem 1. MAX- $r(n)$ -SAT-AA is para-NP-complete for $r(n) = \lceil \log n \rceil$.

Proof. MAX- $r(n)$ -SAT-AA is in para-NP as given a truth assignment for an instance Φ of MAX- $r(n)$ -SAT-AA, we can decide, in polynomial time, whether the assignment satisfies at least $\text{asat}(\Phi) + k$ clauses. To complete our proof of para-NP-completeness, we give a reduction from LINEAR-3-SAT to MAX- $r(n)$ -SAT-AA with $k = 2$.

Consider a 3-SAT formula F with n variables, x_1, \dots, x_n , and m distinct clauses c_1, \dots, c_m . Since F is an input to LINEAR-3-SAT, we may assume that $m \leq cn$ for some positive constant c .

We form a MAX- $r(n)$ -SAT-AA instance F' with $n' = 2cn$ variables, the existing variables x_1, \dots, x_n , together with new variables $y_1, \dots, y_{n'-n}$, and $m' = 2^{\lceil \log n' \rceil + 1}$ clauses. The set of clauses of F' consists of three sets, C_1 , C_2 and C_3 , described below:

- C_1 is the complete set of clauses on variables $y_1, \dots, y_{\lceil \log n' \rceil}$ without the clause consisting of all negative literals, $\bar{c} = (\bar{y}_1 \vee \bar{y}_2 \vee \dots \vee \bar{y}_{\lceil \log n' \rceil})$.
- $C_2 = \{c_i \vee \bar{y}_4 \vee \dots \vee \bar{y}_{\lceil \log n' \rceil} : i \in [m]\}$.
- C_3 is a set of $m' - |C_1| - |C_2|$ clauses on the variables $y_{\lceil \log n' \rceil + 1}, \dots, y_{n'-n}$, of length $\lceil \log n' \rceil$ such that each variable appears in at least one clause and every clause consists of only positive literals.

We claim that F is satisfiable if and only if F' is a YES-instance of MAX- $r(n)$ -SAT-AA for $k = 2$, thus completing the proof. Since, in F' , each clause is of length $\lceil \log n' \rceil$, we have $\text{asat}(F') = (1 - 2/m')m' = m' - 2$. Thus, MAX- $r(n)$ -SAT-AA for $k = 2$ asks whether all the clauses of F' can be satisfied.

Suppose F is satisfied by a truth assignment x^0 . Extend this assignment to the variables of F' by assigning all y_i to be TRUE. Since F is satisfied, all the clauses in C_2 are satisfied. Every clause in C_1 and C_3 contains at least one positive literal, and so is satisfied. Hence F' is satisfied.

If F' is satisfied, then $y_1, \dots, y_{\lceil \log n' \rceil}$ must all be set to TRUE (otherwise, there is a clause in C_1 that is not satisfied). As a result, the set C_2 of clauses can be simplified to the 3-SAT formula F , and thus F must be satisfied. \square

It is not hard to prove Theorem 1 without starting from LINEAR-3-SAT. We use LINEAR-3-SAT to ensure that $n' = O(n)$ which is necessary in the proof of the next theorem. Hereafter, we will assume the Exponential Time Hypothesis (ETH), which is stated below.

Exponential Time Hypothesis (ETH) [16]: There is a positive real s such that 3-SAT cannot be solved in time $2^{sn}n^{O(1)}$. Here n is the number of variables.

Using the sparsification lemma [16, Corollary 1], one may assume that in the input formula F to 3-SAT, every variable occurs in at most p clauses for some positive constant p . For completeness we sketch the proof here.

Lemma 1 ([16]). (Sparsification Lemma) *For every $\varepsilon > 0$ and positive integer r , there is a constant C so that any r -CNF formula F with n variables, can be expressed as $F = \bigvee_{i=1}^t Y_i$, where $t \leq 2^{\varepsilon n}$ and each Y_i is an r -CNF formula with at most Cn clauses. Moreover, this disjunction can be computed by an algorithm running in time $2^{\varepsilon n}n^{O(1)}$.*

Proposition 1 ([16]). *Assuming the ETH, there is a positive real s' such that Linear 3-SAT cannot be solved in time $O(2^{s'n})$.*

Proof. Suppose the proposition does not hold. That is, for all positive real s' LINEAR 3-SAT can be solved in time $O(2^{s'n})$. Now consider a 3-CNF formula F . Using the algorithm for LINEAR 3-SAT we will show that for every positive real c , 3-SAT can be solved in time $2^{cn}n^{O(1)}$, contradicting the ETH. Let $\varepsilon' < c/2$. Using Sparsification Lemma, we produce linear 3-CNF formulas $Y_1, \dots, Y_t, t \leq 2^{\varepsilon'n}$, in time $2^{\varepsilon'n}n^{O(1)}$. We can solve all of them in time $2^{2\varepsilon'n}n^{O(1)} = 2^{cn}n^{O(1)}$ and so obtain a solution for F in time $O(2^{cn})$, a contradiction. \square

Thus, it shows that LINEAR-3-SAT cannot be solved in time $2^{o(n)}$ unless ETH fails. Using the ETH, we strengthen Theorem 1.

Theorem 2. *Assuming the ETH, MAX- $r(n)$ -SAT-AA is not in XP for any $r(n) \geq \log \log n + \phi(n)$, where $\phi(n)$ is any unbounded strictly increasing function of n .*

Proof. Let $\phi(n)$ be an unbounded strictly increasing function of n . Note that if the theorem holds for some unbounded strictly increasing function $\psi(n)$ of n , it holds also for any strictly increasing function $\psi'(n)$ such that $\psi'(n) \geq \psi(n)$ for every $n \geq 1$. Thus, we may assume that $\phi(n) \leq \log \log n$.

Consider a 3-SAT instance I with a linear number of clauses, and reduce it to a MAX- $r(n')$ -SAT-AA instance with $r(n') = \lceil \log n' \rceil$ as in the proof of Theorem 1. Note that I has $O(n')$ variables. Let F' be the formula of the MAX- $r(n')$ -SAT-AA instance and let n be the maximum integer such that $\log n' \geq \log \log n + \phi(n)$. Add $n - n'$ new variables to F' together with a pair of contradicting unit clauses, (x) , (\bar{x}) for each new variable. Let F denote the resulting formula. The total number n of variables in F is such that $r(n) = \lceil \log n' \rceil \geq \log \log n + \phi(n)$. Note that

$$n \leq 2^{n'/2^{\phi(n)}} \leq 2^{n'/2^{\phi(n')}} = 2^{o(n')},$$

and hence it takes $2^{o(n')}$ time to construct F .

Observe that for any value of k , F' is a YES-instance of MAX- $r(n')$ -SAT-AA if and only if F is a YES-instance of MAX- $r(n)$ -SAT-AA. We established in the proof of Theorem 1 that F' is a YES-instance of MAX- $r(n')$ -SAT-AA for $k = 2$ if and only if the 3-SAT instance is satisfiable. Thus, F is a YES-instance of MAX- $r(n)$ -SAT-AA for $k = 2$ if and only if the 3-SAT instance I is satisfiable. Therefore, if there was an XP algorithm for MAX- $r(n)$ -SAT-AA, then for $k = 2$ it would have running time $n^{O(1)} = 2^{o(n')}$, contradicting the ETH. \square

3 Algorithmic Results

To prove the main result of this section, Theorem 4, we reduce MAX- $r(n)$ -SAT-AA to MAX- $r(n)$ -LIN2-AA defined below.

In the problem MAXLIN2, we are given a system S consisting of m equations in n variables, where each equation is of the form $\prod_{i \in I} x_i = b$, $b \in \{-1, 1\}$ and each variable x_i may only take a value from $\{-1, 1\}$. Each equation is assigned a positive integral weight and we wish to find an assignment of values to the variables in order to maximize the total weight of satisfied equations.

Let W be the sum of the weights of all equations in S and let $\text{sat}(S)$ be the maximum total weight of equations that can be satisfied simultaneously. Note that

$W/2$ is a tight lower bound on $\text{sat}(S)$. Indeed, consider choosing assignments to the variables uniformly at random. Then $W/2$ is the expected weight of satisfied equations (as the probability of each equation to be satisfied is $1/2$) and so is a lower bound; to see the tightness consider a system consisting of pairs of equations of the form $\prod_{i \in I} x_i = -1$, $\prod_{i \in I} x_i = 1$ of weight 1, for some non-empty $I \subseteq [n]$. This leads to the following parameterized problem:

MAX- $r(n)$ -LIN2-AA

Instance: A system S of equations $\prod_{i \in I_j} x_i = b_j$, where $b_j \in \{-1, 1\}$, $|I_j| \leq r(n)$, $x_i \in \{-1, 1\}$, and $j \in [m]$, in which Equation j is assigned a positive integral weight w_j , and a nonnegative integer k . Let $W = \sum_{j=1}^m w_j$.

Parameter: k .

Question: Decide whether $\text{sat}(S) \geq W/2 + k$.

The *excess* for $x^0 = (x_1^0, \dots, x_n^0) \in \{-1, 1\}^n$ over S is

$$\varepsilon_S(x^0) = \frac{1}{2} \left[\sum_{j=1}^m c_j \prod_{i \in I_j} x_i^0 \right],$$

where $c_j = w_j b_j$. Observe that $\varepsilon_S(x^0)$ is the difference between the total weight of equations satisfied by x^0 and $W/2$. Thus, the answer to MAXLIN2-AA is YES if and only if $\varepsilon_S(x^0) \geq k$ for some x^0 .

Consider two reduction rules for MAXLIN2 introduced in [14].

Reduction Rule 1. Let A be the matrix over \mathbb{F}_2 corresponding to the set of equations in S , such that $a_{ji} = 1$ if variable x_i appears in equation e_j , and 0 otherwise. Let $t = \text{rank} A$ and suppose columns a^{i_1}, \dots, a^{i_t} of A are linearly independent. Then delete all variables not in $\{x_{i_1}, \dots, x_{i_t}\}$ from the equations of S .

Reduction Rule 2. If we have, for a subset I of $[n]$, an equation $\prod_{i \in I} x_i = b'_I$ with weight w'_I , and an equation $\prod_{i \in I} x_i = b''_I$ with weight w''_I , then we replace this pair by one of these equations with weight $w'_I + w''_I$ if $b'_I = b''_I$ and, otherwise, by the equation whose weight is bigger, modifying its new weight to be the difference of the two old ones. If the resulting weight is 0, we delete the equation from the system.

The two reduction rules are of interest due to the following:

Lemma 2. [14] Let S' be obtained from S by Rule 1 or 2. Then the maximum excess of S' is equal to the maximum excess of S . Moreover, S' can be obtained from S in time polynomial in n and m .

Using techniques from linear algebra, the authors of [6] showed the following:

Theorem 3. Let J be an instance of MAX- $r(n)$ -LIN2-AA in which system S is reduced with respect to Rules 1 and 2. If $n \geq (2k - 1)r(n) + 1$ then the answer to J is YES.

Let I be an instance of MAXSAT-AA given by a CNF formula F with clauses c_1, \dots, c_m , and variables x_1, \dots, x_n . It will be convenient for us to denote TRUE and

FALSE by -1 and 1 , respectively. For a truth assignment $x^0 \in \{-1, 1\}^n$, the *excess* $\varepsilon_I(x^0)$ for x^0 is the number of clauses satisfied by x^0 minus $\text{asat}(x^0)$. Thus, the answer to I is YES if and only if there is an assignment x^0 with $\varepsilon_I(x^0) \geq k$.

MAX- $r(n)$ -SAT-AA is related to MAX- $r(n)$ -LIN2-AA as follows. Results similar to Lemma 3 have been proved in [1, 5].

Lemma 3. *Let I be an instance of MAX- $r(n)$ -SAT-AA with n variables, m clauses and parameter k . Then in time $2^{r(n)}m^{O(1)}$ we can produce an instance J of MAX- $r(n)$ -LIN2-AA with parameter $k2^{r(n)-1}$ such that I is a YES-instance if and only if J is a YES-instance and J is reduced by Rule 2. Moreover, for any truth assignment $x \in \{-1, 1\}^n$, $\varepsilon_J(x) = \varepsilon_I(x) \cdot 2^{r(n)-1}$.*

Proof. Let I be an instance of MAX- $r(n)$ -SAT-AA with clauses c_1, \dots, c_m and variables x_1, \dots, x_n . For a clause c_j , $\text{var}(c_j)$ will denote the set of variables in c_j and r_j the number of literals in c_j . For every $j \in [m]$, let

$$h_j(x) = 2^{r(n)-r_j} [1 - \prod_{x_i \in \text{var}(c_j)} (1 + d_{ij}x_i)],$$

where $d_{ij} = 1$ if x_i is in c_j and $d_{ij} = -1$ if \bar{x}_i is in c_j .

Let $H(x) = \sum_{j=1}^m h_j(x)$. We will prove that for a truth assignment $x \in \{-1, 1\}^n$, we have

$$H(x) = 2^{r(n)} \varepsilon_I(x). \quad (1)$$

Let $q_j = 1$ if c_j is satisfied by x and $q_j = 0$, otherwise. Observe that $h_j(x)/(2^{r(n)-r_j})$ equals $1 - 2^{r_j}$ if $q_j = 0$ and 1 , otherwise. Thus,

$$\begin{aligned} H(x) &= \sum_{j=1}^m [2^{r(n)-r_j} q_j + (2^{r(n)-r_j} - 2^{r(n)})(1 - q_j)] \\ &= 2^{r(n)} [\sum_{j=1}^m q_j - \sum_{j=1}^m (1 - 2^{-r_j})] \\ &= 2^{r(n)} \varepsilon_I(x). \end{aligned}$$

It follows from (1) that the answer to I is YES if and only if there is a truth assignment x such that

$$H(x) \geq k2^{r(n)}. \quad (2)$$

Algebraic simplification of $H(x)$ will lead us to the following:

$$H(x) = \sum_{S \in \mathcal{F}} c_S \prod_{i \in S} x_i, \quad (3)$$

where $\mathcal{F} = \{\emptyset \neq S \subseteq [n] : c_S \neq 0, |S| \leq r(n)\}$. The simplification can be done in time $2^{r(n)}m^{O(1)}$.

Observe that by replacing each term $c_S \prod_{i \in S} x_i$ with the equation $\prod_{i \in S} x_i = 1$ if $c_S \geq 0$ and $\prod_{i \in S} x_i = -1$ if $c_S < 0$, with weight $|c_S|$, the sum $\sum_{S \in \mathcal{F}} c_S \prod_{i \in S} x_i$ can be viewed as twice the excess of an instance J of MAX- $r(n)$ -LIN2-AA. Let $k2^{r(n)-1}$ be the parameter of J . Then, by (2), I and J are equivalent.

Note that the algebraic simplification of $H(x)$ ensures that J is reduced by Rule 2. This completes the proof. \square

It is important to note that the resulting instance J of MAX- $r(n)$ -LIN2-AA is not necessarily reduced under Rule 1 and, thus, reduction of J by Rule 1 may result in less than n variables.

From Theorem 3 and Lemma 3 we have the following fixed-parameter-tractability result for MAX- $r(n)$ -SAT-AA.

Theorem 4. MAX- $r(n)$ -SAT-AA is (i) in XP for $r(n) \leq \log \log n - \log \log \log n$ and (ii) fixed-parameter tractable for $r(n) \leq \log \log n - \log \log \log n - \phi(n)$, for any unbounded strictly increasing function $\phi(n)$.

Proof. We start by proving Part (ii). Let $\phi(n)$ be an unbounded strictly increasing function of positive integral argument. Note that $\phi(n)$ can be extended to a continuous positive strictly increasing function $\phi(t)$ of real argument $t \geq 1$. Thus, $\phi(t)$ has an inverse function $\phi^{-1}(t)$. We may assume that $\phi(n) \geq 0$ for each $n \geq 1$ as otherwise we may consider only n large enough.

Let $r(n) \leq \log \log n - \log \log \log n - \phi(n)$ and consider an instance I of MAX- $r(n)$ -SAT-AA. Note that $2^{r(n)} \leq n$. Therefore by Lemma 3, in polynomial time we can reduce I into an instance J of MAX- $r(n)$ -LIN2-AA, such that I is a YES-instance if and only if J is a YES-instance with parameter $k \cdot 2^{r(n)-1}$. Consider the MAX- $r(n)$ -LIN2-AA instance J' with n' variables formed by reducing J by Rule 1. If $n' \leq \log n$, J' may be solved in polynomial time by trying all $2^{n'} \leq n$ assignments to the variables of J' . Thus, we may assume that $n' > \log n$.

If $n' \geq (k2^{r(n)} - 1)r(n) + 1$, then by Theorem 3 and Lemma 3, I is a YES-instance. Thus, it remains to consider the case $\log n < n' \leq (k2^{r(n)} - 1)r(n)$. We have

$$\log n \leq (k2^{r(n)} - 1)r(n) \text{ and so } \log n \leq k(\log \log n) \cdot \log n / (2^{\phi(n)} \log \log n).$$

This simplifies to $\phi(n) \leq \log k$ and so $n \leq \phi^{-1}(\log k)$. Hence, I can be solved in time $2^{\phi^{-1}(\log k)} m^{O(1)}$ by trying all possible assignments to variables of J' .

Now we will prove Part (i). Let $r(n) \leq \log \log n - \log \log \log n$ and consider an instance I of MAX- $r(n)$ -SAT-AA. As in Part (ii) proof, we reduce I into an instance J of MAX- $r(n)$ -LIN2-AA, such that I is a YES-instance if and only if J is a YES-instance with parameter $k \cdot 2^{r(n)-1}$. Consider the MAX- $r(n)$ -LIN2-AA instance J' with n' variables formed by reducing J by Rule 1. If $n' \leq k \log n$, J' may be solved in XP time by trying all $2^{n'} \leq n^k$ assignments to the variables of J' . Thus, we may assume that $n' > k \log n$. If $n' \geq (k2^{r(n)} - 1)r(n) + 1$, then by Theorem 3 and Lemma 3, I is a YES-instance. Thus, it remains to consider the case $k \log n < n' \leq (k2^{r(n)} - 1)r(n)$. We have $k \log n \leq kr(n)2^{r(n)}$ and so $\log \log n \leq \log \log n - \log \log \log n$, a contradiction. Thus, this case is impossible and we can solve I in XP time. \square

4 Open Problems

It would be interesting to close the gap between the inequalities of Theorems 2 and 4.

Apart from MAXLIN-AA and MAXSAT-AA mentioned above, there are some other constraint satisfaction problems parameterized above a tight lower bound whose complexity has been established in the last few years. One example is r -LINEAR-ORDERING-AA for $r = 2$ and 3. Let $r \geq 2$ be a fixed integer. In r -LINEAR-ORDERING,

given a positive integer n and a multiset \mathcal{C} of r -tuples of distinct elements from $[n]$, we wish to find a permutation $\pi : [n] \rightarrow [n]$ which maximizes that the number of *satisfied* r -tuples, i.e., r -tuples (i_1, i_2, \dots, i_r) such that $\pi(i_1) < \pi(i_2) < \dots < \pi(i_r)$. Let m stand for the number of r -tuples in \mathcal{C} .

Let $\tau : [n] \rightarrow [n]$ be a random permutation (chosen uniformly from the set of all permutations). Observe that the probability that an r -tuple is satisfied by τ is $1/r!$. Thus, by linearity of expectation, the expected number of r -tuples satisfied by τ is $m/r!$. Using conditional expectation derandomization method [2], it is not difficult to obtain a polynomial time algorithm for finding a permutation π which satisfies at least $m/r!$ r -tuples. Thus, we can easily obtain an $1/r!$ -approximation algorithm for r -LINEAR ORDERING. It is remarkable that for any positive ε there exists no polynomial $(1/r! + \varepsilon)$ -approximation algorithm provided the Unique Games Conjecture (UGC) of Khot [17] holds. This result was proved by Guruswami *et al.* [11] for $r = 2$, Charikar *et al.* [3] for $r = 3$ and, finally, by Guruswami *et al.* [10] for any r .

Observe that every permutation π satisfies exactly one r -tuple in the set $\{(i_1, i_2, \dots, i_r) : \{i_1, i_2, \dots, i_r\} = [r]\}$ and, thus, $m/r!$ is a tight lower bound on the maximum number of r -tuples that can be satisfied by a permutation π . It is natural to ask what is the parameterized complexity of the following problem r -LINEAR-ORDERING-AA: decide whether there is a permutation π which satisfies at least $m/r! + k$ r -tuples, where $k \geq 0$ is the parameter. Gutin *et al.* [14] and [12] proved that r -LINEAR-ORDERING-AA is fixed-parameter tractable for $r = 2$ and $r = 3$, respectively. The complexity of r -LINEAR-ORDERING-AA for $r \geq 4$ remains an open problem [12]. Note that if r -LINEAR-ORDERING-AA is fixed-parameter tractable for some r , then all permutation constraint satisfaction problems of arity r parameterized above average are fixed-parameter tractable too (see [12] for the definition of a permutation constraint satisfaction problem of arity r and a proof of the above-mentioned fact).

Acknowledgment This research was partially supported by an International Joint grant of Royal Society.

References

- [1] N. Alon, G. Gutin, E. J. Kim, S. Szeider, and A. Yeo, Solving MAX- r -SAT above a tight lower bound. *Algorithmica*, 61 (2011), 638–655..
- [2] N. Alon and J. Spencer, *The Probabilistic Method*, 2nd Edition, Wiley, 2000.
- [3] M. Charikar, V. Guruswami, and R. Manokaran, Every permutation CSP of arity 3 is approximation resistant. *Proc. Computational Complexity 2009*, 62–73.
- [4] R. Crowston, G. Gutin, M. Jones and A. Yeo, A New Lower Bound on the Maximum Number of Satisfied Clauses in Max-SAT and Its Algorithmic Applications. *Algorithmica*, doi:10.1007/s00453-011-9550-1.
- [5] R. Crowston, G. Gutin, M. Jones, E. J. Kim, and I. Ruzsa. Systems of linear equations over \mathbb{F}_2 and problems parameterized above average. *Proc. SWAT 2010*, Lect. Notes Comput. Sci. 6139 (2010), 164–175.

- [6] R. Crowston, M. Fellows, G. Gutin, M. Jones, F. Rosamond, S. Thomassé and A. Yeo. Simultaneously Satisfying Linear Equations Over \mathbb{F}_2 : MaxLin2 and Max- r -Lin2 Parameterized Above Average. *Proc. FSTTCS 2011*, LIPICS Vol. 13, 229–240.
- [7] M. Cygan, M. Pilipczuk, M. Pilipczuk and J. Wojtaszczyk. On Multiway Cut parameterized above lower bounds. *Proc. IPEC 2011*, to appear.
- [8] R. G. Downey and M. R. Fellows. Parameterized Complexity, Springer, 1999.
- [9] J. Flum and M. Grohe. Parameterized Complexity Theory, Springer, 2006.
- [10] V. Guruswami, J. Håstad, R. Manokaran, P. Raghavendra, and M. Charikar, Beating the random ordering is hard: Every ordering CSP is approximation resistant. *SIAM J. Comput.* 40(3) (2011), 878–914.
- [11] V. Guruswami, R. Manokaran, and P. Raghavendra, Beating the random ordering is hard: Inapproximability of maximum acyclic subgraph. *Proc. FOCS 2008*, 573–582.
- [12] G. Gutin, L. van Iersel, M. Mnich, and A. Yeo, Every ternary permutation constraint satisfaction problem parameterized above average has a kernel with a quadratic number of variables, *J. Comput. System Sci.* 78 (2012), 151–163.
- [13] G. Gutin, M. Jones and A. Yeo, A New Bound for 3-Satisfiable MaxSat and its Algorithmic Application. *Proc. FCT 2011*, *Lect. Notes Comput. Sci.* 6914 (2011), 138–147.
- [14] G. Gutin, E. J. Kim, S. Szeider, and A. Yeo. A probabilistic approach to problems parameterized above or below tight bounds. *J. Comput. Sys. Sci.* 77 (2011), 422–429.
- [15] R. Impagliazzo and R. Paturi. On the complexity of k -SAT. *J. Comput. Sys. Sci.* 62 (2001), 367–375.
- [16] R. Impagliazzo, R. Paturi and F. Zane. Which problems have strongly exponential complexity? *J. Comput. Sys. Sci.* 63 (2001), 512–530.
- [17] S. Khot, On the power of unique 2-prover 1-round games. *Proc. STOC 2002*, 767–775.
- [18] M. Mahajan and V. Raman. Parameterizing above guaranteed values: MaxSat and MaxCut. *J. Algorithms* 31(2) (1999), 335–354.
- [19] M. Mahajan, V. Raman, and S. Sikdar. Parameterizing above or below guaranteed values. *J. Comput. Sys. Sci.* 75(2):137–153, 2009. A preliminary version in the 2nd IWPEC, *Lect. Notes Comput. Sci.* 4169 (2006), 38–49.
- [20] R. Niedermeier. Invitation to Fixed-Parameter Algorithms. Oxford University Press, 2006.
- [21] V. Raman, M. S. Ramanujan and S. Saurabh. Paths, Flowers and Vertex Cover. *Proc. ESA 2011*, *Lect. Notes Comput. Sci.* 6942 (2011), 382–393.

- [22] C.A. Tovey, A simplified satisfiability problem, *Discr. Appl. Math.* 8 (1984), 85–89.